

## [Letter to the editor] Telomere fusions associate with coding sequence and copy number alterations in CLL

Article (Published Version)

Escudero, Laura, Cleal, Kez, Ashelford, Kevin, Fegan, Chris, Pepper, Chris, Liddiard, Kate and Baird, Duncan M (2019) [Letter to the editor] Telomere fusions associate with coding sequence and copy number alterations in CLL. *Leukemia*. ISSN 0887-6924

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/82254/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



Chronic lymphocytic leukemia

# Telomere fusions associate with coding sequence and copy number alterations in CLL

Laura Escudero<sup>1</sup> · Kez Cleal<sup>1</sup> · Kevin Ashelford<sup>1</sup> · Chris Fegan<sup>1</sup> · Chris Pepper<sup>2</sup> · Kate Liddiard<sup>1</sup> · Duncan M. Baird<sup>1</sup>

Received: 6 September 2018 / Revised: 19 October 2018 / Accepted: 11 February 2019  
© The Author(s) 2019. This article is published with open access

## To the Editor:

Short-dysfunctional telomeres are detected prior to clinical progression in chronic lymphocytic leukaemia (CLL) and result in chromosomal fusions that propagate genome instability, driving disease progression. To investigate the impact of telomere dysfunction on the CLL genome, we performed a large-scale molecular characterisation of telomere fusion events in CLL B-cells. A cohort of 276 CLL patient samples was selected for analysis based on short telomere length (TL) profiles, with the majority (97%,  $n = 269$ ) having mean TL within the previously-defined fusogenic range in CLL [1]. Patient samples were screened for the presence of telomere fusions using a single-molecule telomere fusion assay [2] modified to include the 5p telomere (Supplementary Figure 1). Telomere fusions were detected in 72% (198/276) of the samples, which were subsequently arbitrarily stratified by fusion frequency (Supplementary Table 1). Fusions were detected for all telomeres assayed, including the 5p telomere, for which fusions were present in 23% (40/177) of patient samples (Supplementary Figure 2, Supplementary Table 2).

High-resolution characterisation of single-molecule amplified telomere fusions from nine CLL patients with the highest fusion frequency was performed by Illumina

HiSeq4000 paired-end sequencing. Following a customised bioinformatics analysis pipeline [3] and manual curation (Supplementary Figure 3), 914 unique telomere fusions were resolved (Supplementary Figure 4), of which 19% (172/914) involved the 5p telomere (Supplementary Figure 5). Intra-(sister-chromatid) or inter-chromosomal telomere fusion events were identified, as well as recombinations involving non-telomeric loci incorporated into telomere fusions (Fig. 1; Supplementary Figure 4, Supplementary Table 3). These captured loci included; the ancestral telomere at Chr2q13-14 ( $n = 11$ ), mitochondrial DNA ( $n = 4$ ) and other non-telomeric genomic loci ( $n = 78$ ). Complex inter-chromosomal events involving multiple disparate loci were also detected ( $n = 7$ ) (Supplementary Figures 5–9).

Distinct signatures of telomere fusions across the genome could be described for each CLL patient sample (Fig. 1). Two patients (CLL3 and CLL6) displayed simple signatures, defined by the presence solely of intra-chromosomal and/or inter-chromosomal telomere-telomere fusions. In contrast, the CLL8 sample telomere fusion profile revealed abundant genomic linkages, including with the ancestral telomere at 2q13 and mitochondrial DNA. Samples CLL1, CLL2, CLL4, CLL5, CLL7 and CLL9 were characterised by complex signatures with a combination of most or all categories of telomere fusion events identified in this study (Fig. 1; Supplementary Table 4).

Telomere dysfunction is associated with increased genomic instability and disease progression in CLL [1, 4], therefore a comprehensive analysis of all patient-derived telomere fusions with non-telomeric genomic loci was undertaken. Locations and junction sequences pertaining to all 93 (10% total fusions) identified inter-chromosomal fusions were investigated to determine commonality of global or local sequence context as well as providing evidence for the engagement of specific DNA repair processes. These inter-chromosomal genomic fusions were less abundant than pure telomeric inter-chromosomal fusions that represented 38% of all fusions characterised.

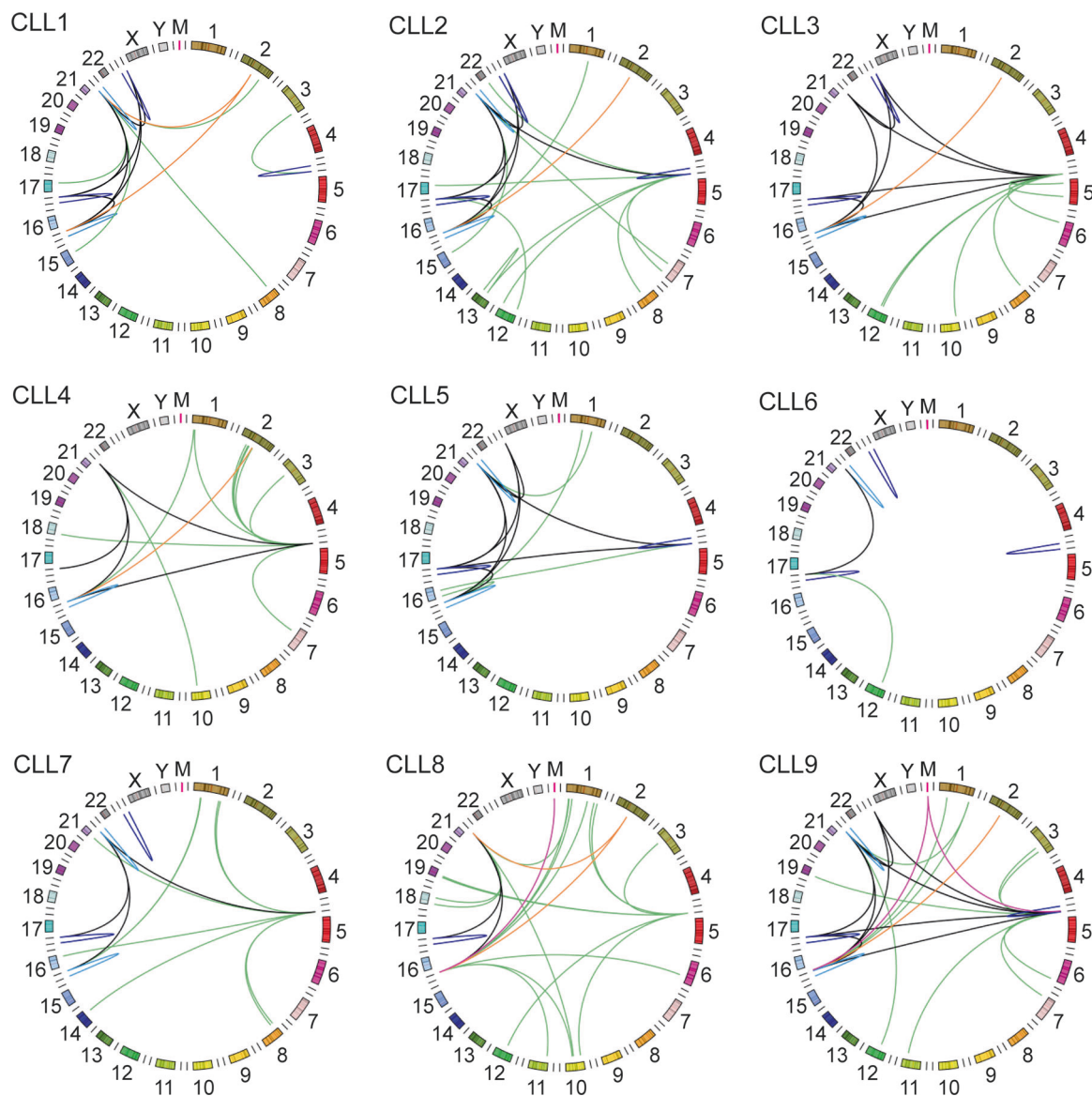
These authors contributed equally: Kate Liddiard, Duncan M. Baird.

**Supplementary information** The online version of this article (<https://doi.org/10.1038/s41375-019-0423-y>) contains supplementary material, which is available to authorized users.

✉ Duncan M. Baird  
[bairddm@cf.ac.uk](mailto:bairddm@cf.ac.uk)

<sup>1</sup> Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK

<sup>2</sup> Brighton and Sussex Medical School, Sussex University, Brighton, UK

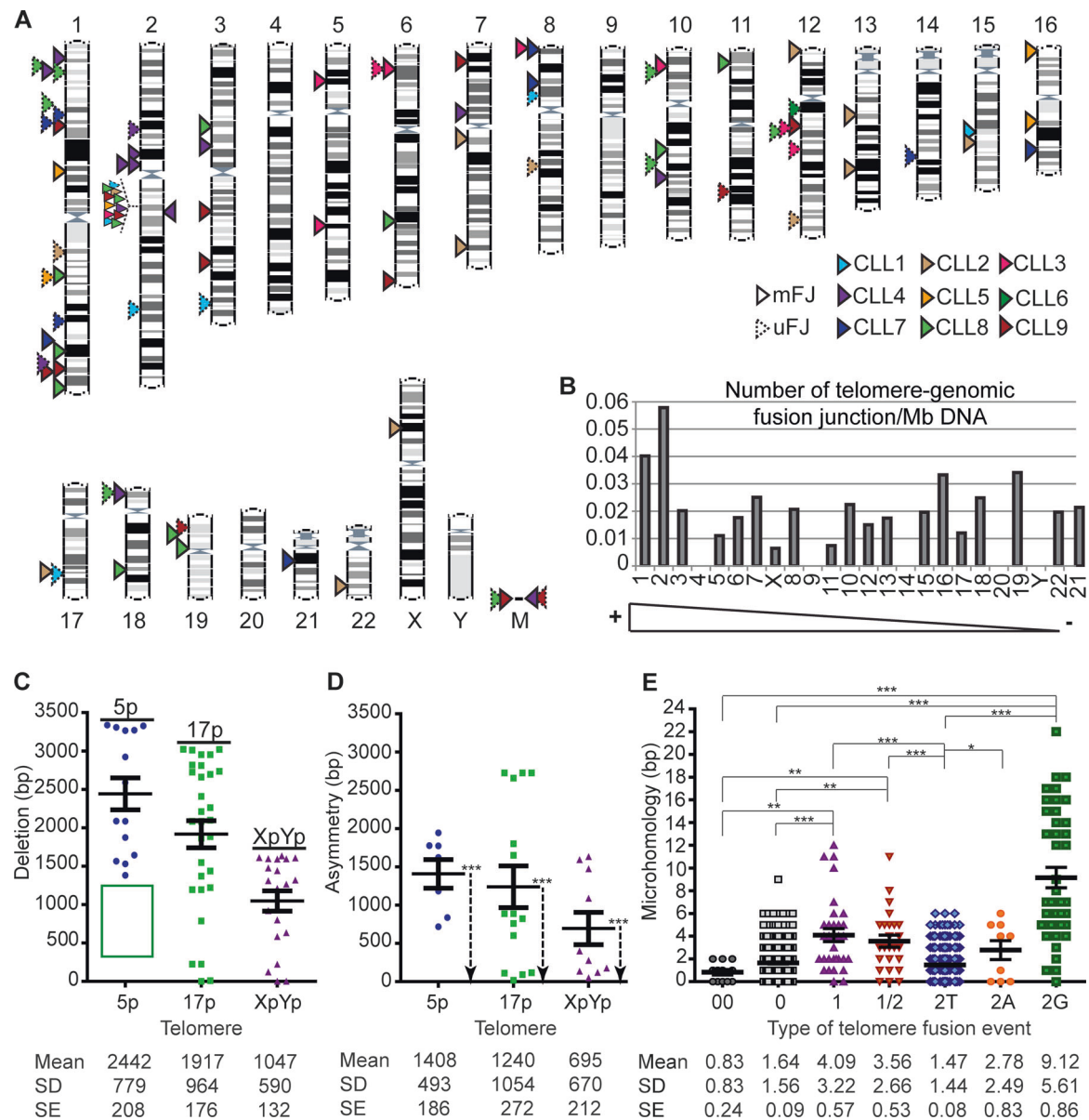


**Fig. 1** Signature of telomere fusions for 9 CLL patient samples. Circos plots showing the validated results obtained from the inter-chromosomal and intra-chromosomal telomere fusion analysis from nine CLL patient samples. Circos plot with each chromosome and its telomeres (1p telomere, Chr1, 1q telomere) around the circle orientated clockwise. Additional notches indicate linkages specifically aligning with subtelomeric sequence references derived from

Stong et al. [12]. Colour code: telomere-telomere inter-chromosomal (black), telomere-telomere intra-chromosomal for 5p, 17p and XpYp (blue), inter-chromosomal or intra-chromosomal for 16p and 21q families (light blue), and inter-chromosomal telomere-genomic (green), telomere-2q13 (orange) and telomere-ChrM (pink). Telomere fusion events with unknown sub-telomeric sequence were not included

Inter-chromosomal fusions with non-telomeric genomic loci were identified in all nine CLL patient samples. Individual events were validated by manual sequence analysis, revealing 68% (63/93) had fusion junctions covered by junction-spanning sequence read pairs (mFJ) and 32% (30/93) had unmapped junctions (uFJ). Each fusion junction location was depicted on the ideogram in Fig. 2a. Notably, the loci disrupted by telomere fusions (summarised in Supplementary Table 5) were not randomly distributed throughout the genome since there was no simple

correlation with chromosome length ( $r^2 = 0.44$ ) or coding gene density of the respective chromosomes ( $r^2 = 0.32$ ) (Fig. 2b, Supplementary Figure 10). However, loci with previously-reported copy number aberrations in CLL [5] were found to be incorporated into telomere fusions, including 2p15, 2p11.2 (2 events), 2q13 (11 events), 6q22.31, 11q22.2 and 18q21.32 (single events). In addition, a complex telomere fusion was detected involving four distinct loci including 13q14.2 that is frequently deleted in CLL (Supplementary Figure 4B).



**Fig. 2** Characterisation of telomere fusions detected across the genome. **a** Validated inter-chromosomal telomere fusion events ( $n = 93$ ) on a karyotype map generated in Ensembl. Telomere fusions with genomic, ancestral telomere 2q13 and mitochondria DNA/Chr. Each colour represents a different patient sample. Continuous arrow-heads indicate mapped fusion junctions (mFJ) and discontinuous arrow-heads represent unmapped fusion junctions (uFJ, location of the read represented). **b** Number of validated inter-chromosomal telomere-genomic fusion junctions per Mb of DNA for each chromosome ordered by length (size obtained from Ensembl). **c** Sister-chromatid deletion and **d** asymmetry for the 5p, 17p and Xp chromosome ends of intra-chromosomal fusion events. Green box highlights the CpG island

Inter-chromosomal telomere fusions occurred within coding DNA more frequently than expected by chance. Over half (57%) of mFJ were within introns and exons of protein-coding genes (Supplementary Table 5), significantly higher (Chi-squared analysis  $p = 0.0024$ ) than the average 42% gene

on the 5p sub-telomere. Location of the fusion primer indicated, determines the limit of the assay from the telomere. **d** Level of asymmetry was determined by calculating the deletion difference between each chromatid of the same fusion event. **e** Microhomology (bp) at the fusion junction was compared for the distinct type of events: TTAGGG-CCCTAA (00), Sub-telomere-TTAGGG (0), intra-chromosomal (1), intra-chromosomal or inter-chromosomal of 16p-16p and 21q-21q families (1/2), inter-chromosomal telomeric fusion events (2T), inter-chromosomal fusions with the ancestral telomere at 2q13 (2A) and inter-chromosomal fusions with genomic loci (2G). Mean, SD and SE are indicated below

content of the human genome (based on the hg19 RefGene). We also observed 15% and 9% mFJ fused with Common Fragile Sites (CFSs) and Alu elements, respectively; however, these were similar to the proportion of CFSs (15%) and Alu sequences (11%) identified across the human genome [6].



All 31 protein-coding genes disrupted by telomere fusions with mFJ were further investigated for potential association with CLL pathogenesis (Supplementary Table 6). An enrichment in genes overexpressed in CD38<sup>+</sup> patient CLL B-cells was revealed using GSEA Gene Set Enrichment Analysis (GSEA, v5.2) Molecular Signatures Database (MSigDB) [7]. This gene set included *HTR7*, *KIF26B* and *LPBN1* ( $p$ -value  $1.5 \times 10^{-6}$ ; FDR  $q$ -value  $2.7 \times 10^{-2}$ )—genes previously found to be upregulated in CD5<sup>+</sup>/CD19<sup>+</sup>/CD38<sup>+</sup> CLL cells associated with worse patient prognosis, compared with patient-matched CD5<sup>+</sup>/CD19<sup>+</sup>/CD38<sup>+</sup> CLL cells in a panel of six patient samples [8]. Strikingly, 36% (11/31) of all genes disrupted by a telomere fusion event for which the junction could be validated were classified as expressed or associated with B lymphocytes or CLL B-cells. These genes included *CD8A*, *RORA*, *TESPA1*, *DMD*, *NOX5*, *NTF3*, *EVI5* and *FTO* (Supplementary Table 7) with documented pathological relevance. A significant enrichment in genes possessing binding motifs matching the B-cell-expressed homeobox transcription factor, HNF1 $\alpha$  (TCF1) [9], within their promoters was also identified (*DMD*, *RORA*, *NTF3* and *HTR7*;  $p$ -value  $2.51 \times 10^{-5}$ ; FDR  $q$ -value  $1.31 \times 10^{-2}$ ; Supplementary Table 8). Furthermore, a noteworthy association of fusion-disrupted genes with gene sets over-expressed in other types of cancer including breast and liver was also revealed by these analyses.

We have previously shown that intra-chromosomal telomere fusion is accompanied by extensive resection that results in asymmetric deletion of the participating sister-chromatids [3]. To assess whether this was true for CLL B-cells, the extent of DNA end-processing at each sister-chromatid was examined for intra-chromosomal fusions with mFJ. The distance from the start of the telomere repeat sequences to the fusion junction for each of the chromatids involved in the fusion event was determined and the difference calculated to obtain a measure of asymmetry (Fig. 2c, d; Supplementary Tables 9–10). The uneven distribution of fusion junctions across the 5p sub-telomere ( $n = 14$ ) is consistent with the location of a CpG island and suggests that the GC-rich sequence may hamper the detection of 5p fusion events (Supplementary Figure 11). Thus, 5p telomere fusions may be under-represented in the data and may have an even greater impact on CLL disease than presently recognised. In contrast, telomere fusion junctions were effectively captured across the 17p ( $n = 30$ ) and XpYp ( $n = 20$ ) telomeres (Fig. 2c). Asymmetry of sister-chromatids was observed for 5p, 17p and XpYp with a mean of 1408 bp, 1240 bp and 695 bp, respectively (Fig. 2d). The degree of asymmetry was significantly greater than the theoretical value 0 (one sample  $t$ -test,  $p < 0.001$ ). This indicates that fusion occurs between sister-chromatids of different lengths in CLL B-cells, consistent

with our observations in other models [2, 3, 10]. No significant differences were found in the extent of asymmetry between the 5p, 17p and XpYp chromosome ends (Kruskal-Wallis,  $p = 0.1661$ ).

High-resolution analysis of each CLL mFJ was performed to investigate candidate DNA repair mechanisms that may underlie distinct types of telomere fusion events. Insertions of templated, untemplated and/or potential telomere variant repeat sequences were observed at 6% (50/796) of mFJ: 23/50 for Telomere-Sub-telomere, 4/50 for intra-chromosomal, 2/50 for intra/inter, 19/50 for telomeric inter-chromosomal, 1/50 for telomere-Chr2q13 and 1/50 for telomere-ChrM fusions. Insertions ranged from 1–21 nucleotides with a mean of 4.5 nucleotides. In contrast, no insertions were identified at fusions with non-telomeric loci. Statistically-significant differences in the extent of microhomology usage at fusion junctions were determined for the different types of telomere fusion events (Kruskal-Wallis  $p < 0.001$  and Dunn's Multiple Comparison Test) (Fig. 2e; Supplementary Table 11). Inter-chromosomal fusions with non-telomeric loci (mean = 9.1 bp;  $n = 43$ ), together with intra-chromosomal sister-chromatid events (mean = 4.1 bp;  $n = 32$ ), displayed the greatest amounts of junction microhomology. In contrast, very low or an absence of microhomology at the fusion point was observed for inter-chromosomal telomeric fusions (mean = 1.5 bp;  $n = 315$ ), Telomere-Telomere (TTAGGG-CCCTAA; mean = 0.8 bp;  $n = 12$ ) and Telomere-Sub-telomere (mean = 1.6 bp;  $n = 303$ ) subgroups. Long tracts of microhomology of up to 23 bp, were observed at inter-chromosomal fusion junctions with non-telomeric loci (Fig. 2e). When the usage of microhomology was >10 bp, the sequence was enriched for the repeat unit of (AC) <sub>$n$</sub>  (Supplementary Figure 5); 40% (6/15) of events that contained at least (AC)<sub>5</sub> (motif ACACACACAC), consistent with repair utilizing single-stranded annealing [11].

Taken together, our data reveal the impact of short-dysfunctional telomeres on the evolving CLL genome, generating tumour heterogeneity that may affect patient prognosis. We have revealed that dysfunctional telomeres predominantly fuse with protein-coding DNA including genes expressed in CLL B-cells and other tumours. We have also identified complex telomere fusions involving multiple non-telomeric loci across the CLL genome, including those with known copy number aberrations in CLL. Our data implicate diverse DNA repair mechanisms at play in CLL tumour initiation and progression, including C-NHEJ, A-NHEJ and SSA. These repair pathways provide potential therapeutic targets and combinations of therapeutic agents targeting these specific pathway components may effectively sensitise CLL B-cell clones with ongoing telomere dysfunction to improve patient outcomes.

**Acknowledgements** The authors would like to thank the CLL patients. The study was undertaken with the support of the UK CLL Clinical Trials Sub-Group. This work was supported by Cancer Research Wales, Cardiff University and Cancer Research UK (C17199/A18246). Patient material was obtained from the UK CLL Trials Biobank, University of Liverpool, which is funded by Bloodwise.

**Author contributions** LE performed the experimental work, analysed the experimental and bioinformatics data and wrote the manuscript; KC carried out the bioinformatics pipeline; CF provided clinical samples and edited the manuscript; CP provided clinical input and edited the manuscript; KL jointly supervised the study and edited the manuscript; DMB jointly supervised the study and edited the manuscript.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Lin TT, Norris K, Heppel NH, Pratt G, Allan JM, Allsup DJ, et al. Telomere dysfunction accurately predicts clinical outcome in chronic lymphocytic leukaemia, even in patients with early stage disease. *Br J Haematol*. 2014;167:214–23.
- Letsolo BT, Rowson J, Baird DM. Fusion of short telomeres in human cells is characterized by extensive deletion and micro-homology, and can result in complex rearrangements. *Nucleic Acids Res*. 2010;38:1841–52.
- Liddiard K, Ruis B, Takasugi T, Harvey A, Ashelford KE, Hendrickson EA, et al. Sister chromatid telomere fusions, but not NHEJ-mediated inter-chromosomal telomere fusions, occur independently of DNA ligases 3 and 4. *Genome Res*. 2016;26:588–600.
- Lin TT, Letsolo BT, Jones RE, Rowson J, Pratt G, Hewamana S, et al. Telomere dysfunction and fusion during the progression of chronic lymphocytic leukemia: evidence for a telomere crisis. *Blood*. 2010;116:1899–907.
- Salaverria I, Martin-Garcia D, Lopez C, Clot G, Garcia-Aragones M, Navarro A, et al. Detection of chromothripsis-like patterns with a custom array platform for chronic lymphocytic leukemia. *Genes Chromosomes Cancer*. 2015;54:668–80.
- Fungtammasan A, Walsh E, Chiaromonte F, Eckert KA, Makova KD. A genome-wide analysis of common fragile sites: what features determine chromosomal instability in the human genome? *Genome Res*. 2012;22:993–1005.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545–50.
- Pepper C, Ward R, Lin TT, Brennan P, Starczynski J, Musson M, et al. Highly purified CD38 + and CD38- sub-clones derived from the same chronic lymphocytic leukemia patient have distinct gene expression signatures despite their monoclonal origin. *Leukemia*. 2007;21:687–96.
- von Wnuck Lipinski K, Sattler K, Peters S, Weske S, Keul P, Klump H, et al. Hepatocyte nuclear factor 1A Is a cell-intrinsic transcription factor required for B cell differentiation and development in mice. *J Immunol*. 2016;196:1655–65.
- Capper R, Britt-Compton B, Tankimanova M, Rowson J, Letsolo B, Man S, et al. The nature of telomere fusion and a definition of the critical telomere length in human cells. *Genes Dev*. 2007;21:2495–508.
- Bhargava R, Onyango DO, Stark JM. Regulation of single-strand annealing and its role in genome maintenance. *Trends Genet*. 2016;32:566–75.
- Stong N, Deng Z, Gupta R, Hu S, Paul S, Weiner AK, et al. Subtelomeric CTCF and cohesin binding site organization using improved subtelomere assemblies and a novel annotation pipeline. *Genome Res*. 2014;24:1039–50.